JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, D08120, doi:10.1029/2008JD010493, 2009

Observed stratospheric downward reflection and its relation to upward pulses of wave activity

N. Harnik[1]

Received 27 May 2008; revised 30 January 2009; accepted 11 February 2009; published 28 April 2009.

[1] We examine the differences between observed stratospheric vertical wave reflection and wave absorption events, which differ in that the wave-induced deceleration remains confined to upper levels in the former. The two types of events signify two types of stratospheric winter dynamics, associated with different downward coupling to the troposphere. Using time-lagged composites, we find that the main factor influencing which event will occur is the duration, in time, of the upward pulse of wave activity entering the stratosphere from the troposphere. Short pulses accelerate the flow at their trailing edge in the lower stratosphere while they decelerate it at upper levels, resulting in a vertical shear reversal, and corresponding downward reflection, while long pulses continue decelerating the vortex at progressively lower levels. The confinement of deceleration to upper levels for short-wave–forcing pulses is also found in an idealized model of an interaction between a planetary wave and the stratospheric vortex, although some aspects of the wave geometry evolution, and thus vertical reflection, are not captured realistically in the model. The results suggest that the stratospheric influence on the type of wave interaction, in reality, is indirect, through a possible effect on the duration of upward wave fluxes through the tropopause.

Citation: Harnik, N. (2009), Observed stratospheric downward reflection and its relation to upward pulses of wave activity, *J. Geophys. Res.*, *114*, D08120, doi:10.1029/2008JD010493.

1. Introduction

[2] The evolution of the winter stratospheric polar vortex is governed by a complex interplay between planetary wave forcing from below and radiative forcing. Thus the net amount of wave activity entering the stratosphere is a central quantity for the evolution and state of the stratosphere. One of the factors complicating the dynamics is the significant influence of the stratosphere itself on the net amount of wave activity entering it [e.g., *Scott and Polvani*, 2004, 2006]. A conceptually simple way in which the stratosphere can influence the net amount of wave activity entering it is by controlling the amount which is reflected back down to the troposphere. *Perlwitz and Harnik* [2003] (hereafter PH03) showed, from observations, that the stratospheric vortex reflects waves down when the jet closes off in the upper high-latitude stratosphere (the vertical shear becomes negative above 10 hPa). They further showed that the northern hemisphere winter circulation can be characterized as *reflecting* or *absorbing*. To show this they defined a reflection index, U_{2-10}, as the mean vertical shear between 2–10 hPa, averaged over 58–74°N. When this shear is negative, a turning surface for vertical propagation is formed because of the meridional gradient of potential vorticity becoming negative. During *reflective* winters (for which the winter mean $U_{2-10} < 0$), an analysis of the time-lagged singular value decomposition between the troposphere and stratosphere shows a peak correlation of planetary wave 1 amplitudes, when the stratosphere leads by a few days, indicative of downward reflection. On the other hand, during absorptive winters (for which $U_{2-10} > 0$), significant persistent correlations are found in the zonal mean fields when the stratosphere leads the troposphere (*Perlwitz and Harnik* [2004], hereafter PH04).

[3] PH03 found that the winter mean U_{2-10} interchanges irregularly between negative and positive values, with about equal occurrence, and that sudden stratospheric warmings occur predominantly during positive winters. In the 90s for example, U_{2-10} was mainly negative and there was a long pause in the occurrence of sudden warmings. One implication for the downward effect of the stratosphere on the troposphere is that the well known downward progression in time-lagged correlations of the Northern Annular Mode index [*Baldwin and Dunkerton*, 2001] disappears when only reflective years are considered and is much more striking when only absorptive years are chosen [PH04]. This suggests that there are two types of winter dynamics (reflective and absorptive) and it raises the question of what determines which state the stratosphere evolves to in any given winter or time period.

[4] Planetary wave forcing from the troposphere varies on seasonal timescales, with maximum forcing during midwinter, but there is also much variability on shorter timescales of 1–2 weeks. *Harnik and Lindzen* [2001] examined the occurrence of downward reflection in the

[1]Department of Geophysics and Planetary Sciences, Tel Aviv University, Tel Aviv, Israel.

0148-0227/09/2008JD010493$09.00

Southern Hemisphere, and found that during mid-winter, reflection occurs intermittently, as a result of the shorter timescale variations, while in fall, the climatological jet structure is reflecting. The Northern Hemisphere is more like Southern Hemisphere winter, in that reflection occurs because of the short timescale variations of wave forcing, when an anomalously strong wave pulse propagates upward and decelerates the flow only in the upper stratosphere. If the deceleration continues and extends downward, a sudden warming typically ensues. Thus we need to find out what determines whether deceleration remains confined to the upper stratosphere (reflection) or whether it spreads down to lower levels (as in a sudden warming and a downward migration of negative Northern Annular Mode anomalies).

[5] As a clear first step in answering this question, we compare the wave fluxes and vortex evolution during observed Northern Hemisphere winter wave reflection and sudden warming events (the latter representing strong wave absorption events). Section 2 introduces the data, the analysis methods and the definition of reflection and sudden warming events. Section 3 presents the results, starting with examination of the shorter timescale evolution during event peaks (section 3.1), and shows the duration of upward wave pulses plays a central role in determining subsequent dynamics. We then examine observed significant early time-lagged anomalies, to assess the possibility of a vortex preconditioning (section 3.2). In section 4 we examine the role of the duration of upward pulses of wave activity in a simple Holton-Mass type model, and we conclude in section 5.

2. Data and Analysis

[6] We use the 2.5 $\times$ 2.5 degree pressure level ERA40 data, daily averaged from 1 September to 31 March, for the years 1979–2002 (the satellite era). We note that the results we present here also hold when repeated on the entire ERA40 data set (1958–2002). To identify reflection events, we use the index defined by PH03: $U_{2-10} = \langle U_{2\mathrm{hPa}} - U_{10\mathrm{hPa}} \rangle$ where $\langle U \rangle$ is the zonal mean zonal wind, averaged over 58–74°N, with a climatological seasonal cycle removed. PH03 and PH04 found that negative seasonal or monthly U_{2-10} values are a good indication of downward reflection. An examination of U_{2-10} based on shorter timescales reveals that while negative U_{2-10} is a necessary condition for reflection, it is often negative during or just before sudden warming events. Thus, in defining a reflection index based on U_{2-10}, care should be taken to only look at negative U_{2-10} events for which winds in the lower stratosphere remain strong. Thus we first find potential reflection events, then find sudden warming events, and when relevant, remove them from the potential reflection event dates. In what follows, we denote refection events by REF, and sudden warmings by SW.

[7] Potential reflection events are defined as those times on which a 3 day running mean of U_{2-10} is 2 standard deviations below its mean value for more than 3 consecutive days. The mean and standard deviation of U_{2-10} are 7.6 and 10.7 (m/s change over the 8 hPa layer) respectively, so that reflection events are defined when $U_{2-10} < -13.8$ m/s. This value is characteristic of the negative shear found in the reflective state of PH03. Events that are less than 3 days apart are counted as one if U_{2-10} between them is negative. A total of 31 events were found during the time period analyzed (for time-lagged compositing we exclude events with a central date after 16 March). We also examined composites based on lower standard deviation thresholds. The resulting composites, which were based on more events but weaker in amplitude, were qualitatively similar to those presented here, with slightly smaller amplitudes (as expected).

[8] To identify sudden warming (SW) events (which represent wave absorption), we adopt the definition used in the study by *Charlton and Polvani* [2007], to be those days for which the daily U(60°N, 10 hPa) becomes negative, followed by at least 10 consecutive days with positive U values (to distinguish from final warmings). Events which occur less than 20 days after a previous one are considered to be a continuation of the same warming. A total of 13 events were found for the period examined (note that we require 10 recovery days before 31 March). For time-lagged composites, the central date is taken to be the first day on which U(10 hPa, 60°N) becomes negative.

[9] An examination of the SW events reveals that U_{2-10} is generally negative during sudden warmings (there is no reason why U(10 hPa) and U(2 hPa) $-$ U(10 hPa) cannot both be negative at the same time). To insure that we only examine reflection events, we exclude from the potential reflection events found above, those for which a sudden warming occurs within its duration or within 2 days after its end (12 negative U_{2-10} events). This yields a total of 19 reflection (REF) events. Since there is a tendency for REF events to occur in succession, we also defined a subset of REF events, where events occurring less than 20 days after a previous event were discarded (referred to as REF-SEP). This yielded a total of 14 events. As we will show shortly, REF events, and related dynamical processes, are characterized by quite a short timescale (1–2 weeks). Thus there is generally enough temporal separation between reflection events, and for much of the results the composites of REF and REF-SEP are very similar. We will show results from both, as indicated.

[10] The time-lagged composites of reflection events are based on a central event date which is defined as the day on which U_{2-10} is a minimum. Since the events are rather short-lived, the results are not sensitive to the definition of this date. (When a threshold value of 1 standard deviation is used, some events consist of a few consecutive weaker upper level decelerations. For those, we also composited things around the first day on which U_{2-10} decreased below the 1 standard deviation threshold. This yields similar composites, only shifted earlier by 3–4 days.) Anomalies for the composites are defined as the deviations from the climatological seasonal cycle. To remove influences of interannual variability, we further remove the winter mean from each winter's anomaly field. Statistical significance is estimated using a Monte-Carlo approach, as follows. We randomly choose 1000 combinations of N days, N being the number of composite members, calculate the corresponding time-lagged composites, and find the maximum, minimum, and 5th and 95th percentile values. As expected, the resulting random data composite maps are independent of the time lag, and have a spatial structure similar to the standard deviation maps of the composited field. All the results presented here have been verified to be robust to details of the composite calculation and event definitions. This included changing (within

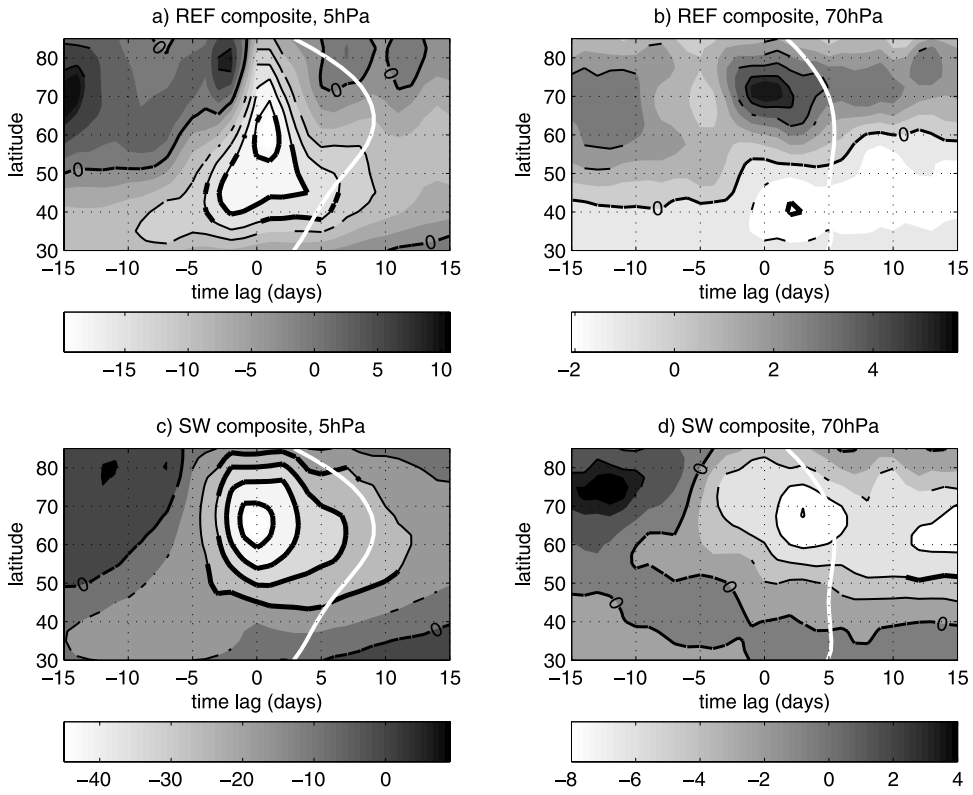

Figure 1. Latitude–time-lagged composite plots (shaded) of zonal mean wind anomaly. (a) REF at 5 hPa. (b) REF at 70 hPa. (c) SW at 5 hPa. (d) SW at 70 hPa. Values are shown by the gray shading, which goes from bright (negative) to dark (positive). Thick and thin contour lines (except for the zero contour, which is marked by a thick line as well) denote values significant at the 99.9% and 95% levels, respectively. Also shown for comparison (white contour) is the climatological wind field for the period in which reflection events occur (December to March), scaled (by the same value at all levels) to fit the plot.

reasonable limits) the way U_{2-10} is defined, the way we exclude reflection events due to a coincidence with a sudden warming and the definition of the seasonal cycle.

[11] To get a sense of the amount of variability which reflection events represent, also in relation to SW events, we perform an empirical orthogonal function (EOF) analysis of the high-latitude vortex (the meridionally averaged zonal mean wind over high latitudes). We find that two EOFs capture most of the vertical variations of the high-latitude vortex (97%). SW events are related to one phase of the leading EOF (which explains 83% of the variance), while REF events are related to one phase of the second EOF (which explains 14% of the variance). Thus these two types of events describe a large part of the high-latitude vortex variability. We note that SW are associated with much more variance than REF events, as expected for deeper structures. It is also possible that some of this difference in variance is due to the different nature of these events – SW are much more violent than reflection, which represents, in some sense, a noninteraction between the waves and the mean flow. The results of this analysis are presented in Appendix A.

3. Results

3.1. Event Center

[12] In this section we examine the differences between reflection and sudden warming events, with the goal of finding the main factors that determine which of these events will occur following a given pulse of wave activity into the stratosphere. Figure 1 shows the polar vortex evolution during the two types of events, as seen in time-lagged composites of U anomalies at 2 levels, in the upper (5 hPa) and lower (70 hPa) stratosphere. During reflection (REF) events, negative U anomalies develop at upper levels around 35°N and spread poleward before strengthening, and eventually decaying (becoming nonsignificant) at around day +3 at high latitudes, and day +7 at lower latitudes. At lower levels, the U anomaly pattern is reversed, with high-latitude positive anomalies below the main deceleration center. Comparing to SW events, those also start with a negative U anomaly at low latitudes which spreads poleward before strengthening, but the anomaly is much stronger, longer lived, and extends deeper than for REF. The significant negative zonal mean wind anomalies throughout and beyond day +15 are consistent with other studies which show very long persistence for SW anomalies [e.g., *Limpasuvan et al.*, 2004]. We also note significant (at the 95% level) positive anomalies at days −15 – −10, in the lower stratosphere, for both types of events (which extend slightly higher for REF events). These early anomalies suggest a vortex preconditioning which will be discussed in greater detail in section 3.2.

[13] Figure 2 shows anomaly composites of the leading terms in the transformed Eularian mean (TEM) version of

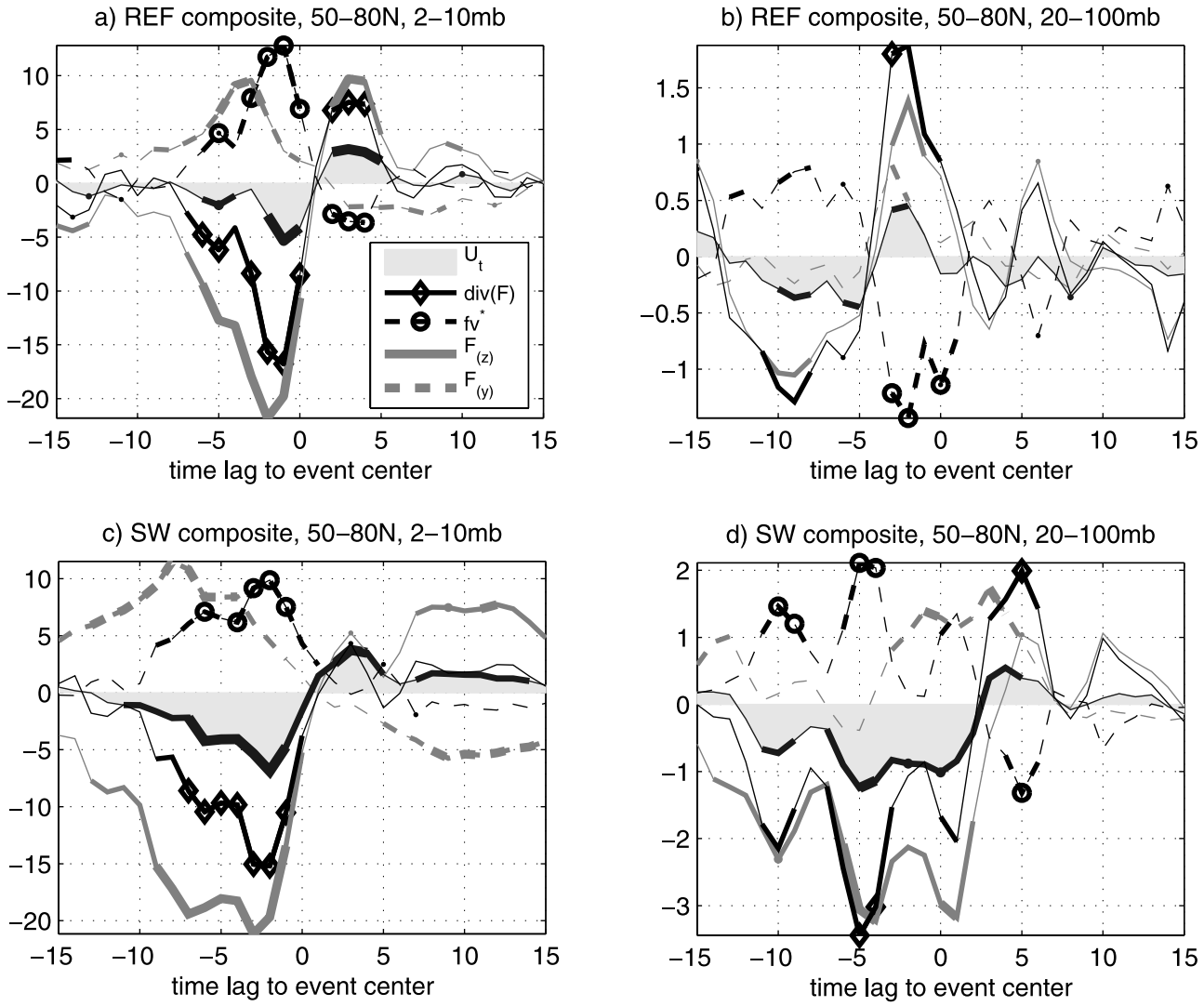

Figure 2. Time-lagged composites of the leading terms in the momentum budget equations, averaged over the latitude range 50–80°N, for the upper stratosphere (2–10 hPa) (a) REF and (c) SW events and lower middle stratosphere (20–100 hPa) (b) REF and (d) SW events. Shown are the deceleration ($U_t(t) = U(t) - U(t-1)$, gray-filled black contour), $\nabla \cdot \mathbf{F}/(\rho a_e \cos(\varphi))$ (solid black diamonds), $f\overline{v}^*$ (dashed black circles), and the vertical (solid gray) and meridional (dashed gray) contributions to $\nabla \cdot \mathbf{F}/(\rho a \cos(\varphi))$. Units are m/s/day. Statistically significant values at the 99.9% and 95% levels are marked using contour thickness. For $\nabla \cdot \mathbf{F}/(\rho a_e \cos(\varphi))$ and $f\overline{v}^*$, we marked the 99.9% by symbols.

the zonal momentum equation (e.g., equations (3.5.1)–(3.5.3) *Andrews et al.* [1987]), box-averaged over the high-latitude upper stratosphere (2–10 hPa), and the middle to lower stratosphere (20–100 hPa). Overall, deceleration is driven by the EP flux divergence ($\nabla \cdot \mathbf{F}$) and opposed by the TEM Coriolis force ($f\overline{v}^*$), as expected from theory [e.g., *Andrews et al.*, 1987]. Consistent with Figure 1, the upper and lower level patterns are opposite each other around the REF event onset date, with lower level anomalies being slightly less significant ($\nabla \cdot \mathbf{F}$ and $f\overline{v}^*$ anomalies are 99.9% significant, while their sum (the acceleration) is significant only at the 95% level). Compared to REF events, the SW upper level deceleration (Figure 2b) is slightly stronger, starts about 5 days earlier relative to event center, and the acceleration which follows is slightly weaker than the deceleration (giving rise to longer lived negative zonal mean wind anomalies seen in Figure 1). Unlike REF events, the lower level U_t pattern is similar to upper levels, only weaker (Figure 2d).

[14] A further decomposition of $\nabla \cdot \mathbf{F}$ into the contributions from the vertical (heat flux) and meridional (momentum flux) components (gray lines) shows a clear dominance of the vertical flux for both types of events, with the meridional component opposing the vertical one, except during REF events in the lower stratosphere, where it precedes the vertical flux by a few days. We note that at high latitudes, the climatological momentum fluxes are diverging. At lower latitudes where momentum fluxes converge, they do drive at least half of the initial deceleration pattern (not shown).

[15] The above REF pattern is suggestive of a short-lived wave forcing, which induces deceleration at its growth

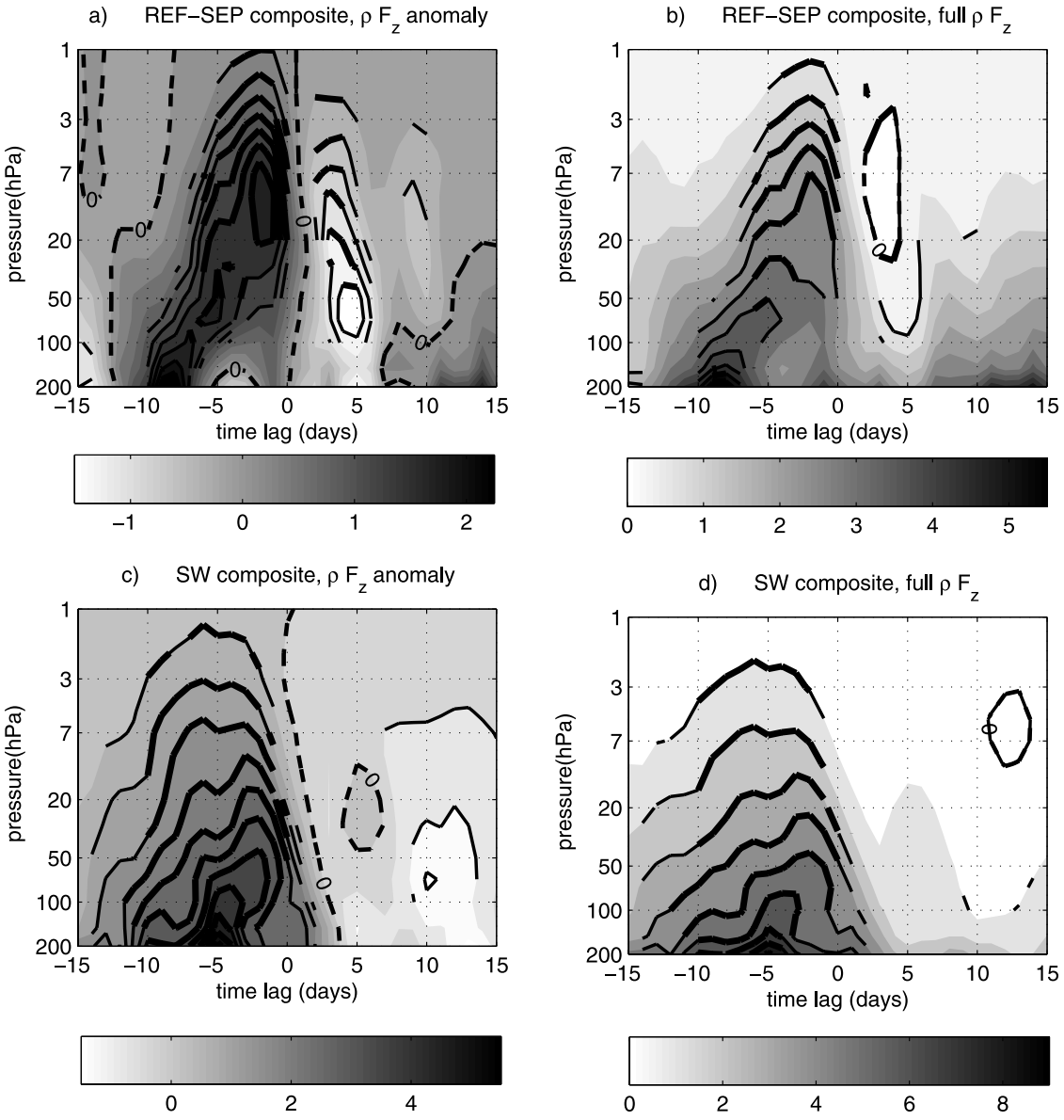

Figure 3. Latitude-height composites of $\tilde{F}_{(z)}(\rho F_{(z)}$ averaged over 50–80N, weighted by cosine latitude). (a) REF-SEP events, $\tilde{F}_{(z)}$ anomalies (deviation from climatology). (b) REF-SEP, total $\tilde{F}_{(z)}$ field. (c) SW, $\tilde{F}_{(z)}$ anomalies. (d) SW total $\tilde{F}_{(z)}$ field. The zero contour is marked by a dashed line. Other contours, in both the anomaly and full field plots, mark regions in which the *anomaly* composites are significant at the 95% (thin) and 99.9% (thick) levels. Note that the vertical axes are stated in pressure (hPa), but their spacing is based on height (log pressure).

stage, where $F_{(z)}$ converges (affecting lower levels first, then upper levels) and acceleration at its decline, where $F_{(z)}$ diverges (day −5 at lower levels, compare equations (3.5.2) and (3.6.3) of *Andrews et al.* [1987]). The SW pattern, on the other hand, is consistent with a longer lived forcing. Indeed, from comparing various other wave and zonal mean flow composites, the most robust and significant difference between the REF and SW events is the relatively short duration of the upward wave activity fluxes prior to REF events. This is shown explicitly in Figure 3, which shows time-lagged height composites of the high-latitude density weighted vertical EP flux ($\tilde{F}_{(z)}$), both anomaly and full fields. To best highlight this result, we present the REF-SEP composite (the REF composite shows the same result very clearly, though not as cleanly). (This result is very robust to all variations we tried, including lowering the U_{2-10} standard deviation threshold from 2 to 1 in the definition of the REF events.) For both events, there is an anomalously strong upward burst of wave activity at negative time lags, followed by a strong weakening of the fluxes (negative anomalies) at day zero. However, $\tilde{F}_{(z)}$ is about twice as strong during SW, and maybe more importantly, its "source", as represented by the 200 hPa values, is much longer lived prior to SW than to REF-SEP events. During REF-SEP events, $\tilde{F}_{(z)}$ starts diminishing from below before peak values are reached at upper levels, resulting in a

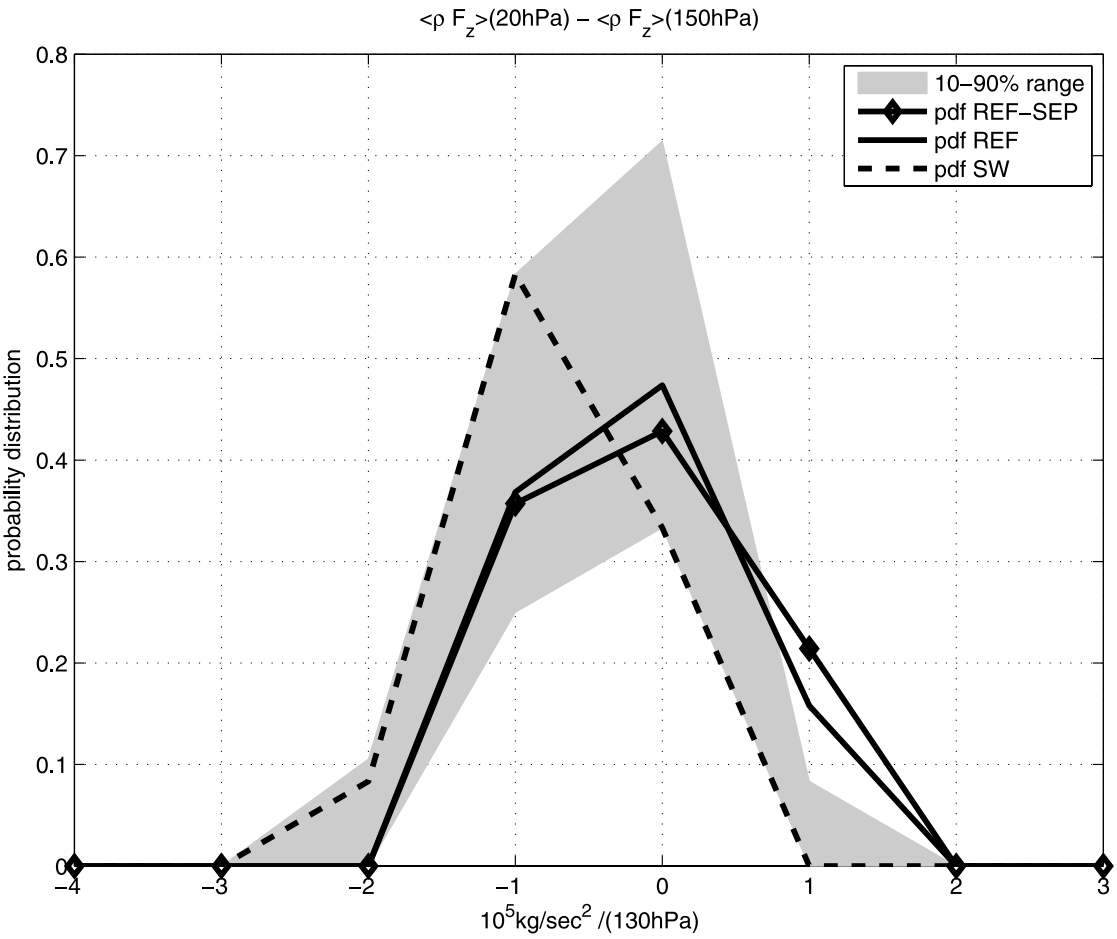

Figure 4. Probability distribution function (histogram divided by total number of events) of $\tilde{F}_{(z)}$, averaged over time lags −4:0 days for REF (solid), REF-SEP (solid diamonds), and SW (dashed). For reference, we also show the 10–90% percentile region (shading), estimated by calculating the pdfs of 1000 random samples of 12, 14, and 19 events (the number of SW, REF-SEP, and REF events, respectively), calculating percentile statistics for each of these samples, and taking the maximal 10–90% range based on the three.

divergence of $\tilde{F}_{(z)}$ below 10 hPa (days −7 to 0). This is consistent with the observed acceleration at lower levels during REF-SEP events, and it promotes the formation of a reflecting surface, by enhancing the formation of negative shear in the upper stratosphere. During SW events, the composite $\tilde{F}_{(z)}$ field is pretty much converging throughout the stratosphere (consistent with Figure 2).

[16] This finding, that reflection forms as a result of the wave pulses being very short, has a few implications. The formation of a reflection surface can be considered as a process which reduces the amount of wave absorption in the stratosphere. However, we now find that reflection only occurs when the wave source itself weakens, thus it only amplifies a reduction in wave absorption which is ultimately caused by a different process. Moreover, if the initial wave pulse which sets up the reflecting surface is short-lived, it is not obvious that there is a remaining wave to be reflected down. We know there must be, however, from PH03, who found a statistically significant downward propagating signal in wave number one during reflective winters (for which U_{2-10} is negative). A clue can be found in Figures 3b and 3d, which shows the full $\tilde{F}_{(z)}$ field composites (as opposed to anomaly composites). The full field consists of a climatological seasonally varying $\tilde{F}_{(z)}$, which is large and positive during mid winter, and the anomalous $\tilde{F}_{(z)}$ (shown in Figures 3a and 3c). We see that the climatological flux is large enough to keep the total fluxes at 200 hPa positive throughout the REF event. It is this climatological flux which is then reflected down by the reflecting surface.

[17] To verify that the differences in $\tilde{F}_{(z)}$ composites are not dominated by a few uncharacteristic extreme events, we further calculate, for each REF, REF-SEP, and SW event, the 10–100 hPa layer mean $\tilde{F}_{(z)}$ divergence (using the full field, rather than anomalies), averaged over the 5 days leading to the event central date, and plot the histogram normalized by the total number of events (which yields a probability distribution function, pdf) in Figure 4. Though on the basis of a small number of events (19, 14 and 12 for REF, REF-SEP and SW respectively), the histograms are nicely distributed around a central value. Moreover, $\tilde{F}_{(z)}$ divergence is generally larger during REF and REF-SEP events (around zero values) and is negative during SW events. To estimate the statistical significance of these differences, we also chose 1000 random sets of event central

dates (chosen from December to March) and calculated the height-time mean $\tilde{F}_{(z)}$ divergences and corresponding pdfs. The gray shading marks the 10th and 90th percentile values of the 1000 pdfs. We see that the $\tilde{F}_{(z)}$ divergence during REF events is indeed significant, while the large $\tilde{F}_{(z)}$ convergence during SW events is only marginally significant (unsurprising, for such a small sample). Nonetheless, the calculation shows that the shorter duration of upward EP flux forcing found in REF composites is not due to a few uncharacteristic extreme events.

3.2. Possible Vortex Preconditioning

[18] In the previous section we examined the most significant features of the composites, which occur around the event central date. We note, however, that there are also signs in the data of significant anomalies at quite long negative time lags. This can be seen in Figure 1 which shows significant (at the 95% level) U anomalies for REF events at time lags -15: -10 days, at 70 hPa, poleward of 50°N, and slightly earlier and at higher latitudes at 5 hPa, as well as significant anomalies for SW during those times at 70 hPa. These anomalies could suggest an indirect stratospheric influence on the occurrence of REF and SW events, through a vortex preconditioning. A possible preconditioning might be due to a tendency of waves to propagate higher up, and interact more strongly with the mean flow (leading either to a REF or a SW event), when the vortex is strong at high latitudes. Significant positive early time anomalies could, on the other hand, simply reflect the fact that by choosing to average over the times when the vortex is weakest, we necessarily remove these states from early times in the composites (though note that the REF composite does not include SW events, and vice verse). In the following section we will examine the early time anomalies, to determine whether they reflect a real tendency, and whether there is a difference at early times between REF and SW events. We note that a strengthening and poleward contraction of the polar vortex prior to sudden warming has been noted and examined in quite a few studies before [e.g., *McIntyre*, 1982; *Limpasuvan et al.*, 2004, and references therein].

[19] An examination of zonal mean wind composites for long negative time lags reveals significant positive wind anomalies as early as -55 days, which weaken slowly (so that the anomalies are still positive for a long time), up to the abrupt REF and SW event weakening. We note that *Limpasuvan et al.* [2005], in their study of vortex intensification events, found a similar behavior of a slow onset and fast decay (opposite to the fast onset and slow decay of negative anomalies during sudden warmings), even though in that study the composites were centered around the peak intensification times, rather than around the peak deceleration times.

[20] To get a better sense of whether the early positive anomalies arise from a few extreme events or from a general tendency of all events, we also calculate the daily histograms (e.g. the histogram of the 19 values at day -5 of the 19 REF events) of vortex strength and upward EP flux anomalies, for the time range -55: 15 days. The resulting pdf–time lag plots (not shown) suggest there are three main periods, a very early period (lags -55: 36 days) where the vortex is clearly anomalously strong, an early period (lags -35: -10 days) where the vortex is still anomalously strong but less so, and the active event period (-5: 0) where the vortex is weak and the upward EP fluxes are anomalously strong. The corresponding pdfs of U and $F_{(z)}$ anomalies, averaged over these three time periods are shown in Figure 5, for REF, REF-SEP and SW events. The vortex strength is measured by the anomalous zonal mean wind, averaged over 30–80°N, 100–5 hPa, and the upward EP flux anomaly is averaged over 30–80°N, and chosen at 30 hPa. (The results at 30 hPa are representative of $F_{(z)}$ at a range of altitudes (about 70–10 hPa), though at early time lags, the anomalies are not as clear lower down.) For reference, we also show the estimated 10th–90th percentile range, based on the pdfs of 1000 randomly chosen sets of event dates (as in Figure 4). The vortex is clearly anomalously strong at very early time lags (-55: -36 days, Figure 5a), more so preceding SW events than REF events. This is true of the vast majority of events, as is clear from the comparison between the composite pdfs and the random pdfs' spread range. We next examine whether the anomalously strong vortex is due to anomalously weak upward fluxes. We find that the $F_{(z)}$ pdfs (plot b), which are very similar for REF and SW events, and are only slightly shifted to anomalously negative values relative to the random sample. At smaller negative time lags (-35: -10), the $F_{(z)}$ pdfs (plot d), which are again very similar for REF and SW. are anomalously strong (the pdf curves are at the upper end of the gray shading for positive values) suggesting the vortex should weaken during this period, and indeed, the wind pdf (plot c) shows that while the vortex is still anomalously strong, it is less so than at earlier time lags. For comparison, we also present the pdfs during the event active period (days -5: 0). As expected, these show very clear anomalies in both U and $F_{(z)}$, with the SW anomalies being stronger than REF for both fields. We note that REF and REF-SEP are very similar for all fields. Also, while there is a clear difference in vortex strength for SW and REF events during all time periods, the differences in the distributions of upward EP fluxes only occur very close to the event central date. Its possible that the vortex structure at early times has an effect on the efficiency of upward propagation of strong wave forcing pulses, however, since both events are characterized by an early anomalously strong vortex, it is hard to pinpoint specific features which lead to subsequent differences.

4. Dependence on the Duration of Upward Wave Pulse in a Holton-Mass Model

[21] In this section, we use an idealized wave–mean flow interaction model to see if we can capture the observed effect of the duration of wave forcing pulses on the corresponding wave-induced deceleration pattern. We use the model of *Holton and Mass* [1976], which is a quasi-geostrophic (QG) β plane channel model of wave–mean flow interaction, in which a single zonal wave number forced from below, propagates vertically on a zonally symmetric mean flow, and modifies it. The mean flow evolves because of the convergence of wave fluxes (only the vertical EP flux component exists in this model), with radiative forcing relaxing the flow back to the initial state. The model is described in Appendix B. Figure 6a shows the

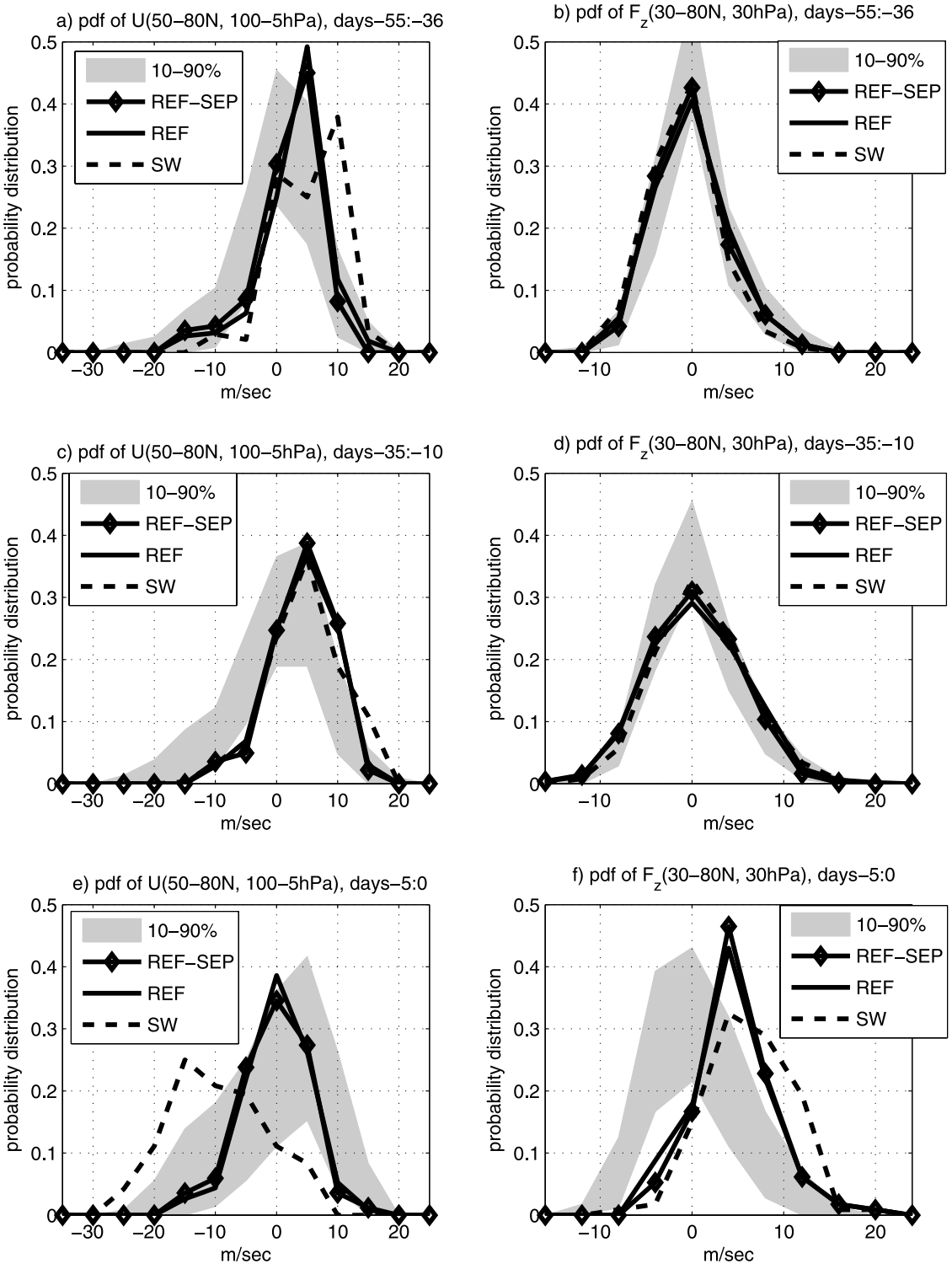

Figure 5. (a, c, e) Pdfs of average high-latitude stratospheric vortex strength anomaly (U(50−80°N, 100−5 hPa) and (b, d, f) extratropical (30−80°N) upward EP flux anomaly at 30 hPa , for REF (solid), REF-SEP (solid diamonds), and SW (dashed) events, calculated daily for each time lag and averaged over three time lag periods: for Figures 5a and 5b, −55:−36 days; Figures 5c and 5d, −35:−10 days; and Figures 5e and 5f, −5:0 days. Also shown are the 10−90% percentile region (shading), estimated as in Figure 4.

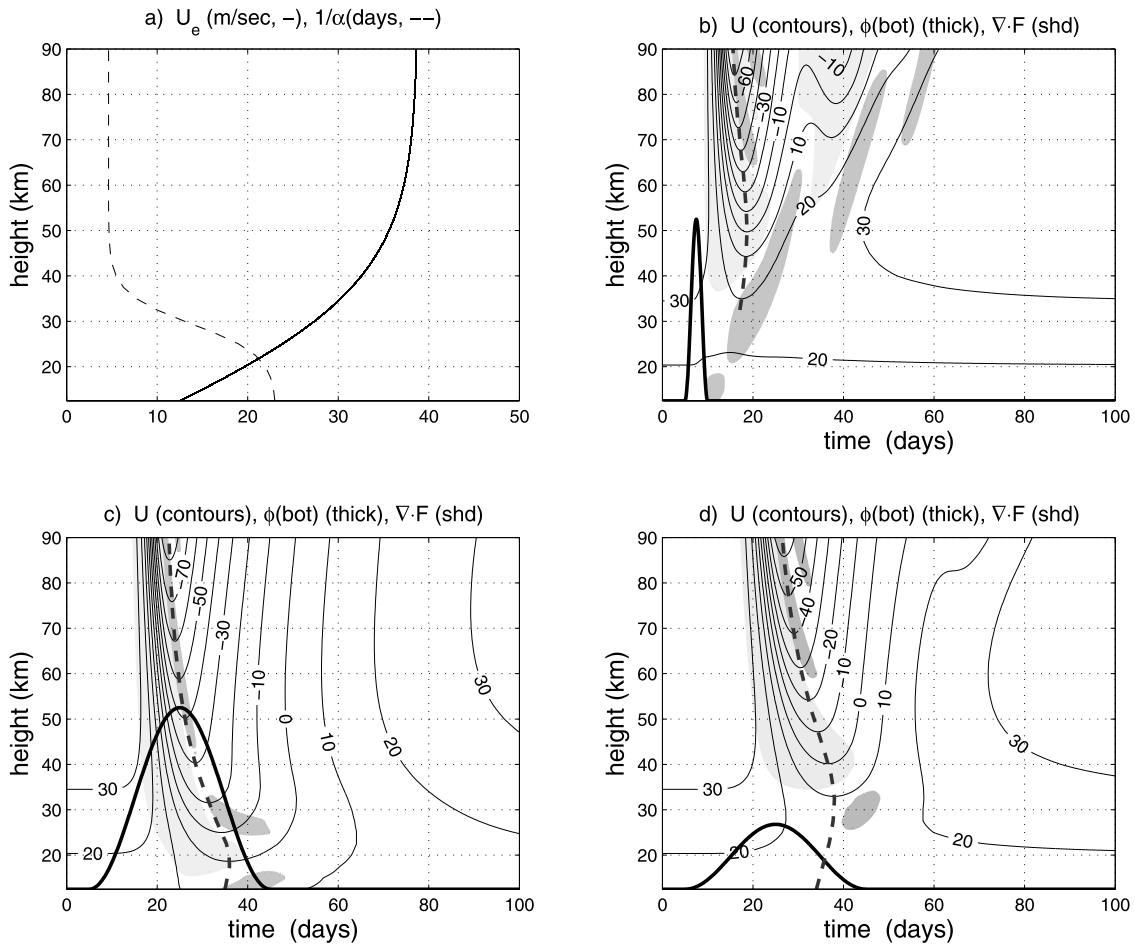

Figure 6. (a) Radiative equilibrium zonal mean wind profile (solid, m/s) and the radiative damping timescale (dashed, *days*) of the Holton-Mass model runs (see text for details). (b) Time-height plot of the zonal mean wind (contours; contour interval, 10 m/s), for the short (5 days) wave pulse, shown in terms of geopotential height amplitude (thick solid line, units of 10 m), shifted by 12.5 (the lowest model level) to fit the plot. Also shown is the line of minimum winds (thick dashed) and $\nabla \cdot F$ (bright and dark gray shading for decelerations stronger than -10 m/s/day and acceleration above 1 m/s/day). (c) Same as in Figure 6b but for a forcing timescale of 40 days, and same maximum amplitude. (d) Same as in Figure 6c but with the time integrated forced wave activity density ($\int \Phi_{\text{bot}}^2 \, dt$) as in Figure 6b.

radiative equilibrium state (U_e) to which we relax the model, along with the relaxation timescale, used both for the mean flow and the waves.

[22] We force the model with a single pulse of stationary wave number 1, by specifying the wave geopotential height at the bottom surface (Φ_{bot}). We vary the timescale of the pulse from 5 to 75 days, once keeping the peak amplitude fixed (run series *R*1), and once keeping the total forced wave activity fixed ($\int |\Phi_{\text{bot}}|^2 dt = const$, run series *R*2). Figures 6b–6d show the time-varying forcing amplitude (thick solid curve) and time-height contours of the zonal mean wind U, for the shortest forcing pulse run which is the same in both run sets (plot b), and the longer forcing pulse runs of *R*1 and *R*2 (plots c, d respectively). Comparing plot *b* to plots *c*, *d* shows the effect of pulse length, while comparing plots *c* and *d* shows the effect of pulse amplitude. The dashed line marks the zonal mean wind minimum. Following this line we see that deceleration starts from the top and propagates downward, and that it lasts longer and reaches lower down, as the pulse length in increased. This is consistent with the observational analysis of the previous section. Also shown, in shading, is the contribution of the vertical EP flux component to the deceleration ($\frac{1}{\rho}\frac{\partial F_{(z)}}{\partial z}$), with the lighter shading denoting deceleration stronger than 10 m/s/day and darker shading denoting acceleration larger than 1 m/s/day. (We note that the deceleration pattern is strongest above about 40 km where the Newtonian damping is strongest. The wave driven acceleration, on the other hand, is mostly due to the transience of the pulse, hence it is much weaker.) We see acceleration just below the peak deceleration, during the short pulse event (plot b), which is consistent with the wave source already starting to shut off by the time the wave packet reaches the upper stratosphere and decelerates the flow there. On the other hand, the longer pulses are still in their growing phase when the upper level deceleration starts, so we do not see a lower level acceleration during the event peak. This supports our observational-based hypothesis that short-wave pulses lead to the formation of a reflecting surface. Comparing Figures 6c and 6d, we

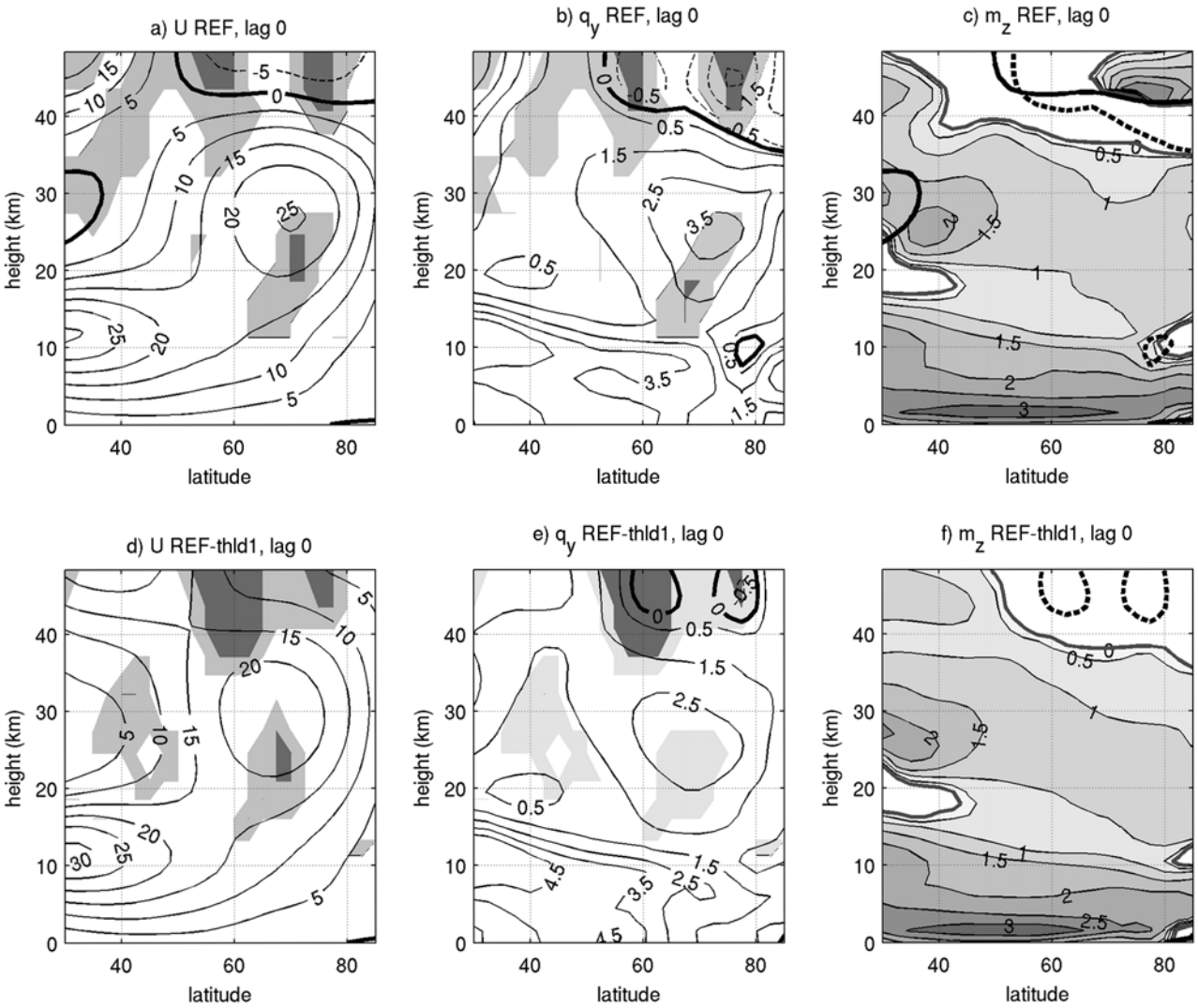

Figure 7. The total field, lag 0, REF composite of (a) U, (b) $\frac{1}{a}\overline{q}_{\varphi}$, and (c) vertical wave number m_z, calculated from the observations. In Figure 7c, only vertical wave propagation regions (real m_z) are shaded, with the vertical turning surface ($m_z = 0$) denoted by a thicker gray line, and zero U and $\overline{q}_y$ surfaces are plotted in thick solid and dashed lines respectively. (d–f) Same as in Figures 7a to 7c but for REF-1SD (events defined based on a 1 standard deviation U_{2-10} index). In Figures 7a, 7b, 7d, and 7e, the light and dark shadings are the 95% and 99.9% significance levels, respectively (calculated from the corresponding anomaly composites); negative values are dashed, and the zero line is thick. Note that only the observations domain is shown, although the model used to calculate $m^2_{(z)}$ extends higher.

see that wave amplitude also affects the level to which downward propagation reaches. This suggests differences in wave forcing amplitude also contribute to the observed differences between REF and SW events (forcing prior to SW is stronger, as seen in Figure 3).

[23] The above runs were done to test the hypothesis that the deceleration remains confined to the upper stratosphere when the forcing pulse is short. There are two parts to our claim: first, when the forcing is short-lived, the deceleration does not have time to reach lower levels. Our model results clearly support this part. The second part suggests that when the forcing is short, there is a region in the stratosphere where the upward EP fluxes are diverging (the pulse trailing edge), resulting in acceleration and a further enhancement of the negative vertical shear. While there is evidence for this in our model, the magnitude of lower level acceleration is weak compared to the upper level deceleration – the lower stratospheric wind anomaly is about $\frac{1}{10}$ the magnitude of the upper stratospheric wind anomaly in the model, compared to about $\frac{1}{3}$ in observations (e.g., Figures 1 and 6). Thus, without examining the wave geometry explicitly, it is not certain that the wave-induced acceleration pattern really leads to wave reflection in our model, during the short pulse runs.

[24] To have a reference, we first examine the observed *wave geometry* of the zonal mean flow during REF events. This was done for the Northern Hemisphere by PH03, but only for seasonal or monthly mean flows. Wave geometry is the characterization of the mean flow in terms of regions of wave propagation/wave evanescence (where the mean flow

supports/does not support vertical or meridional propagation of waves with a specific zonal wave number and phase speed). The main quantity in this context is the index of refraction squared, N^2_{ref}. A positive N^2_{ref} indicates wave propagation, while a negative N^2_{ref} indicates wave evanescence. In the real atmosphere waves propagate both meridionally and vertically. To examine propagation in the vertical direction, we isolate the vertical part of N^2_{ref}, using the *vertical wave number* diagnostic ($m^2_{(z)}$) developed in the study by *Harnik and Lindzen* [2001] (described briefly in Appendix C). In analogy to N^2_{ref}, a positive (negative) vertical wave number squared indicates vertical wave propagation (evanescence). For stationary waves with zero zonal phase speed, the leading term in N^2_{ref} and in $m^2_{(z)}$ is the ratio between the meridional gradient of potential vorticity, $\frac{1}{a}\overline{q}_\varphi$, and the zonal mean wind (see equation (C1)). When this ratio is negative, N^2_{ref} and $m^2_{(z)}$ are generally negative too. Of particular importance are critical surfaces, where N^2_{ref} and $m^2_{(z)}$ become singular due to a vanishing of the zonal mean wind, and turning surfaces, where $m^2_{(z)}$ vanishes, due for example to $\frac{1}{a}\overline{q}_\varphi$ becoming small or negative. Waves will break and get absorbed in the mean flow when reaching a critical surface while they will get reflected at a turning surface.

[25] Figure 7 shows REF composites of the full U and $\frac{1}{a}\overline{q}_\varphi$ fields from observations, with the statistically significant regions shaded (calculated from the corresponding anomaly composites). We see the clear region with negative vertical shear in the upper stratosphere, with a zero wind line forming in the high-latitude stratosphere at around 42 km (Figure 7a), within a region of negative $\frac{1}{a}\overline{q}_\varphi$ which extends slightly below the zero wind line poleward of 50°N (Figure 7b). For both fields, the zero lines occur within the statistical significance regions. The corresponding $m^2_{(z)}$ composite shows a clear turning surface forming at around 35 km (Figure 7c, gray line at the edge of the shaded region). Also shown in Figures 7d, 7e, and 7f are the composites of REF events based on a lower U_{2-10} threshold of one standard deviation (REF-1SD, we note that all the results shown for REF events also hold for REF-1SD). We see again a negative $\frac{1}{a}\overline{q}_\varphi$ region forming (Figure 7d), along with a turning surface (Figure 7f), but with no critical surface, indicating this is only a feature of the stronger deceleration events. We note that this configuration of a critical surface forming above a turning surface will theoretically over-reflect waves approaching it from below [*Lindzen*, 1988; *Harnik and Heifetz*, 2007].

[26] We now turn back to the Holton-Mass model, and examine it for evidence of downward reflection. Figure 8a shows the time integrated upward EP flux, $F_{(z)}$, divided by the time integrated theoretical value of $F_{(zz)}$, for an upward propagating pure plane wave, as a function of pulse length, for the two run sets. The theoretical value is calculated assuming the mean flow is constant with height using the values at the lowest level. Since the actual EP flux is a combination of upward and downward propagating components (in a WKB sense), we expect the time integral of this quantity to be small for downward reflected waves, and on the order of 1 for upward propagating waves. We note that our forcing, via wave amplitude, does not set the upward EP flux, which depends on the vertical phase tilt. Rather, this is determined as part of the solution, and is affected by the forcing and the mean flow state in a complex, nonlinear manner [*Esler and Scott*, 2005]. Indeed, this ratio is smallest for short pulses, increases quite rapidly with pulse duration from 0.4 to around 1.2, and it saturates for pulses longer than about 20 days. In addition, the forcing amplitude does not affect this ratio too much (compare $R1$ and $R2$). This suggests that indeed, for short enough pulses, we get downward reflection, which results in a reduction of the net amount of wave activity entering the stratosphere.

[27] We next examine the mean flow wave geometry and wave structure during the short-pulse model runs, for evidence of downward reflection. Since our model is in Cartesian coordinates, and is separable in latitude and height, the expression for index of refraction squared is slightly different than for the observations, and is denoted by $\tilde{N}^2_r$. The exact expression and its relation to N^2_{ref} is described in Appendix C. We use the vertical profile of wave geopotential height phase, as an indication of the wave propagation direction, with a decrease of wave phase with height indicating net downward propagation. For clarity, we only show the wave phase in regions where the wave amplitude exceeds a certain value. Figure 8b shows U and $\overline{q}_y$ at two times, before and during the peak deceleration, for the short pulse run of Figure 6b, while Figure 8c shows the corresponding $\tilde{N}^2_r$ and wave phase structure during these times. The dashed lines represent the initial state, while the solid lines represent the peak of the deceleration event. We see that both U and $\overline{q}_y$ become negative in the upper stratosphere during the deceleration event, but unlike the observations, U becomes negative lower down than $\overline{q}_y$. This means that a critical surface forms below a turning surface, evident as a positive-below-negative spike dipole in $\tilde{N}^2_r$ at around 47–50 km (note that the critical surface falls between model grid points so that $\tilde{N}^2_r$ is finite). Examining the wave phase structure, we see that initially the wave phase increases with height throughout the domain, consistent with upward wave propagation. During the peak deceleration, however, the phase decreases with height in a small region below 30 km. Since there is no turning surface in this run, this decrease with height could indicate partial reflection from the minimum $\tilde{N}^2_r$ region at around 60 km. Indeed, this phase decrease with height only occurs when this minimum in $\tilde{N}^2_r$ forms. This is consistent with the smaller EP flux ratio of Figure 8a, but this ratio might be due to some complex wave–mean flow evolution near the critical surface, rather than simple downward reflection.

[28] To try and force the model to behave more like the observations, where a turning surface forms (with the zero wind line forming above the zero $\overline{q}_y$ line), we decrease the wave forcing by half, and increase the radiative equilibrium wind profile by 10 m/s, calling this run set $R3$. We expect this to raise the zero wind line. (We note that *Giannitsis and Lindzen* [2009], in an idealized wave–mean flow model of the stratosphere, found that the mean flow evolution, in particular, whether a critical surface or turning surface forms, is sensitive to the magnitude of the initial zonal mean wind.) The results are shown in plots d, e. We see that although the zero wind line has moved up by about 20 km, the zero $\overline{q}_y$ line moved up too, so that now they are roughly the same height. Nonetheless, this is sufficient to form a reflecting surface, where $\tilde{N}^2_r = 0$, at around 60 km (note that $\tilde{N}^2_r$ can becomes zero when $\overline{q}_y$ is small, provided U is large

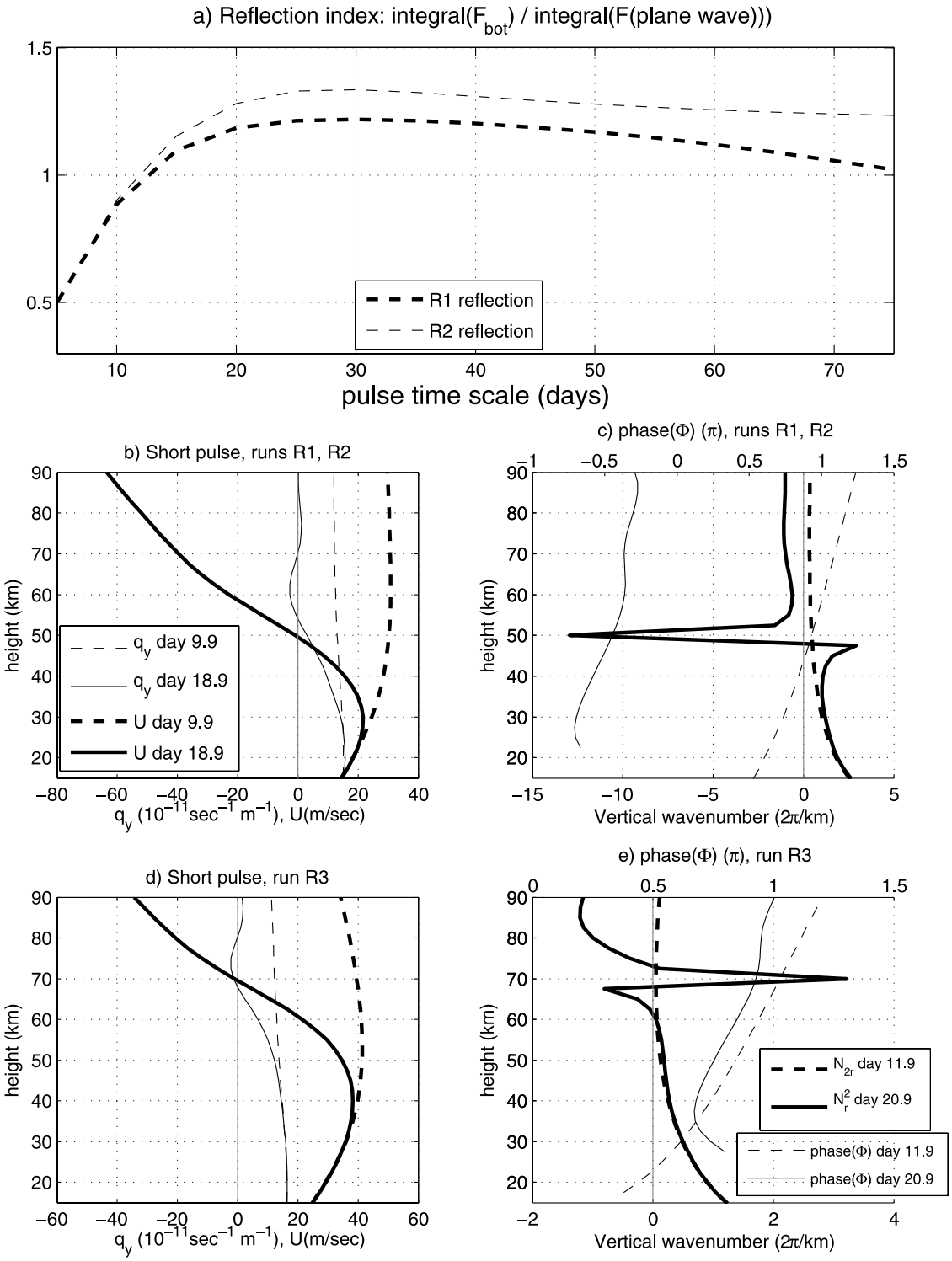

Figure 8. (a) Ratio between the actual and theoretical pure-plane wave time integrated upward EP fluxes through the bottom level (as a measure of wave reflection, see text for details), for run sets *R*1 (thick lines) and *R*2 (thin lines). (b) *U* (thick lines) and $\overline{q}_y$ (thin lines) before (day 9.9, dashed) and during the peak deceleration (day 18.9, solid) for the 5-day pulse run of *R*1. (c) $\tilde{N}_r^2$ (equation (C4), thick lines) and wave geopotential height phase (thin lines). (d, e) As in Figures 8b and 8c but for a weaker wave forcing and stronger radiative equilibrium wind (run *R*3), before (day 11.9, dashed) and during the peak deceleration (day 20.9, solid). The wave phase lines (in Figures 8c and 8e) are only shown for regions where the wave amplitude exceeds 100 m.

enough). Correspondingly, the decrease of wave phase with height becomes more pronounced (plot e).

[29] The above results suggest that wave reflection does occur for a short-wave–forcing pulse in the Holton-Mass model, though in a more limited way than in the observations. The ratio of lower stratospheric acceleration to upper stratospheric deceleration is weaker in the model than in observations. Consistently, a zero wind line always forms near the region of negative $\overline{q}_y$, but whether it forms above or below it depends on the model parameters. While in observations the effect of q_y dominates, in the model U tends to dominate. Two main shortcomings of the simple model which might cause this difference are the idealized thermal damping, and the lack of meridional propagation. Thermal damping influences the vortex both by directly accelerating it, and by affecting eddy fluxes. The meridional convergence of EP fluxes, while not being the leading term driving the observed high-latitude deceleration/acceleration pattern, does affect it significantly. In particular, it contributes significantly, and positively to the observed acceleration at lower levels (e.g. dashed gray line in Figure 2, in particular plot b). This contribution which enhances lower level acceleration, is completely lacking in our model. A weak acceleration results is a smaller decrease in $\overline{q}_y$, and thus in $\tilde{N}_r^2$.

5. Discussion

[30] In this work we examine the differences between two observed types of stratospheric wave-mean flow interaction events. In the first type of event (REF) the upward propagating wave decelerates the vortex at upper levels only, which results in the formation of a reflecting surface and subsequent downward reflection of the remaining wave activity. In the second type of event (SW), the waves decelerate the vortex throughout the stratosphere, leading to a major warming. As shown by PH04, the two types of events signify two types of stratospheric winter dynamics, associated with different downward coupling to the troposphere. This suggests a way in which the stratosphere can influence the troposphere, by influencing which type of wave–mean flow dynamics evolves. It also opens way for a possible positive feedback, if for example, a weak polar vortex state favors SW events which further weaken it, and a strong polar vortex favors REF events, which result in less wave absorption.

[31] To examine the validity of this hypothesis for observations, we compare the mean flow, wave fluxes, and various terms in the momentum budget of these two types of events, using composite analysis. Rather than support the hypothesis above, our findings indicate that a main factor influencing which event will occur is the duration, in time, of the upward pulse of wave activity entering the stratosphere from the troposphere. Short pulses, which shut off, or significantly weaken, before or just as they start decelerating the flow at upper levels, also accelerate the flow at their trailing edge in the lower stratosphere. As a result the vertical shear of the mean zonal wind becomes negative, and a reflecting surface forms which reflects the rest of the wave pulse back down. On the other hand, long pulses, which last longer than it takes the waves to reach upper levels and decelerate the vortex, continue decelerating it at progressively lower levels, raising the chances for the occurrence of a sudden warming. Thus long upward pulses seem to be a necessary ingredient for a SW to occur. A similar result was found by *Zhou et al.* [2002], when they examined what caused some polar warming events to propagate downward and some not. (Though *Zhou et al.* [2002] were not thinking of downward reflection, a comparison of their events with ours suggests there is a connection; out of their 7 downward propagating warming events, 6 coincide with a SW and 1 with a REF event, while of their 7 non propagating events, 5 coincide with REF events, and one encompasses both a SW and REF event.) *Chen and Robinson* [1991, 1992] found that shorter pulses are actually more efficient in ''sending'' wave activity upward into the stratosphere, due to the projection onto higher frequencies (a linear effect). Our results suggest that the wave–mean flow interaction makes the absorption of such waves less efficient.

[32] Our results emphasize the importance of determining the source of variations in upward EP flux through the troposphere-stratosphere boundary. While this is still an open question, several studies have indicated that the stratosphere itself influences the amount of wave activity entering it from below [e.g., *Scott and Polvani*, 2004, 2006; *Chen and Robinson*, 1992; *Esler and Scott*, 2005], suggesting an indirect way by which the stratosphere can influence the troposphere. An examination of long time-lagged composites of zonal mean wind do indeed show persistent and significant positive high-latitude U anomalies during the month or two prior to REF events, however, these positive anomalies are also as strong prior to SW events. Thus our observational analysis did not indicate any clear way in which the stratospheric state can affect the upward pulses of wave activity.

[33] An idealized model of a single zonal wave number, interacting with a vertically varying mean flow, is able to capture the confinement of deceleration to the upper stratosphere for short-wave pulses. However, the simple model does not capture the wave geometry evolution very realistically. In particular, unlike the observations, the model is not able to give pure reflection, in which $\tilde{N}_r^2$ becomes negative without a zero wind line forming. This may possibly be due to the lack of meridional wave propagation in the simple model. In the observations, meridional fluxes and meridional wave breaking act to reduce upper level deceleration, while they accelerate the flow in the lower stratosphere. We note that *Giannitsis and Lindzen* [2009] found that wave-wave interactions were important for capturing the mean flow evolution, but that given the mean flow evolution, linear wave dynamics were able to capture the full nonlinear wave evolution well.

[34] To summarize, this study suggests that downward reflection does not *initiate* a reduction in the amount of wave activity being absorbed in the stratosphere. Rather, it reflects wave activity back down during times when its source from the troposphere is greatly reduced because of another process. Thus it is more of a ''symptom'' than an active player. Nonetheless, as was shown statistically by PH03 and PH04, downward reflection of waves influence the tropospheric circulation by changing the amplitude and phase of wave 1, and moreover, during reflective winters, the zonal mean downward influence is shut off. Thus it is

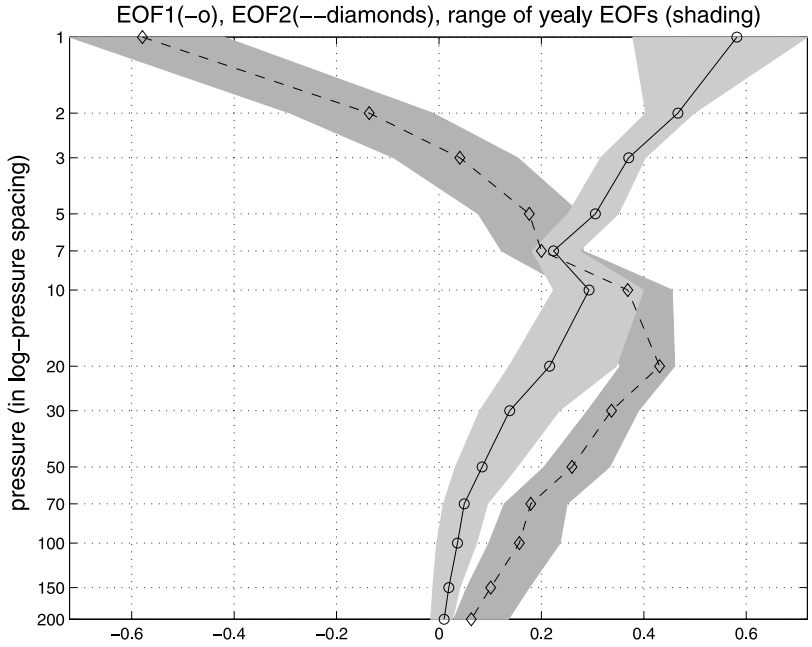

Figure A1. EOFs of the high-latitude (58–74°N) averaged zonal mean wind, using daily data for 1 September to 31 March (January, years of 1979–2002) between 200 and 1 hPa (13 levels, shown by grid lines). Shown are the first (solid, circles) and second (dashed, diamonds) EOFs calculated using all years of data. Shading indicates the range of values spanned by the EOFs calculated based on individual winter data. The EOFs are calculated using an area weighting (cos latitude, and the height spacing between vertical grid points).

still important to understand when downward reflection occurs, and to study its direct and indirect influence on the troposphere, both on synoptic and longer timescales.

Appendix A: The High-Latitude Vortex Variability and Its Relation to REF and SW Events

[35] To get a sense of the relation of the variability of the zonal mean wind structure, and the occurrence of reflective states (negative U_{2-10}) we calculate a high-latitude EOF based on the zonal mean winds averaged over 58–74°N (latitudes used for the definition of U_{2-10}), for all observational levels between 200–1 hPa (13 levels), using the extended-winter daily data described above. We use a cosine latitude weighting, as well as a vertical grid height weighting (to account for the uneven spacing of the observations with height). We do not weight the winds by pressure, since we are interested in the wind structure, rather than in the momentum or acceleration fields. Figure A1 shows the vertical structure of the first two EOFs, which explain 0.83 and 0.14 of the total high-latitude variance, respectively. The first EOF represents a weakening-strengthening of the polar vortex at all levels, with largest amplitude at upper levels (which is consistent with wave deceleration being stronger higher up due to the density effect). (The kink at 7 hPa is due to the vertical grid space weighting (it is not found when such weighting is not done).) The second EOF peaks between 20–30 hPa, and has a node between 5–3 hPa so that a negative anomaly at upper levels corresponds to a positive anomaly in the lower-mid stratosphere. We define the EOFs such that they are positive when the zonal mean wind at 10 hPa is positive, which means that an upper stratospheric deceleration corresponds to a negative PC1 and a positive PC2.

[36] The vertical structures of the EOFs suggests sudden warmings correspond to an anomalously negative PC1, (*Limpasuvan et al.* [2004] used an EOF of U(50 hPa) as a definition of sudden stratospheric warming events. They found 39 events (during 1958–2001), only 26 of which were categorized as major sudden warmings, based on the WMO definition.) while REF events correspond to an anomalously positive PC2. A closer look at the PC time series along with the REF and SW event dates (Figure A2) does indeed support this. We see that the vast majority of REF events coincide with positive peaks of PC2 (Figure A2, blue curves and diamonds), while SW events tend to coincide with negative peaks of PC1 (Figure A2, black curves and stars).

[37] We calculate the correlations between the two PC time series and a few wind indices: U_{2-10} (for REF events), U(10 hPa, 60°N) and U(58–74, 10 hPa) (for sudden warmings). The results, shown in Table A1, indicate a high correlation between PC1 and the sudden warming indices (0.73 and 0.94) and between PC2 and the reflection index (−0.79). These correlations are significant even if we assume only one degree of freedom per year. We also verify that the high correlations come from intraseasonal, rather than interannual, timescales, by calculating the yearly EOFs based on individual winters. We find a very similar structure of the first two EOFs, with PC1 accounting for most of the

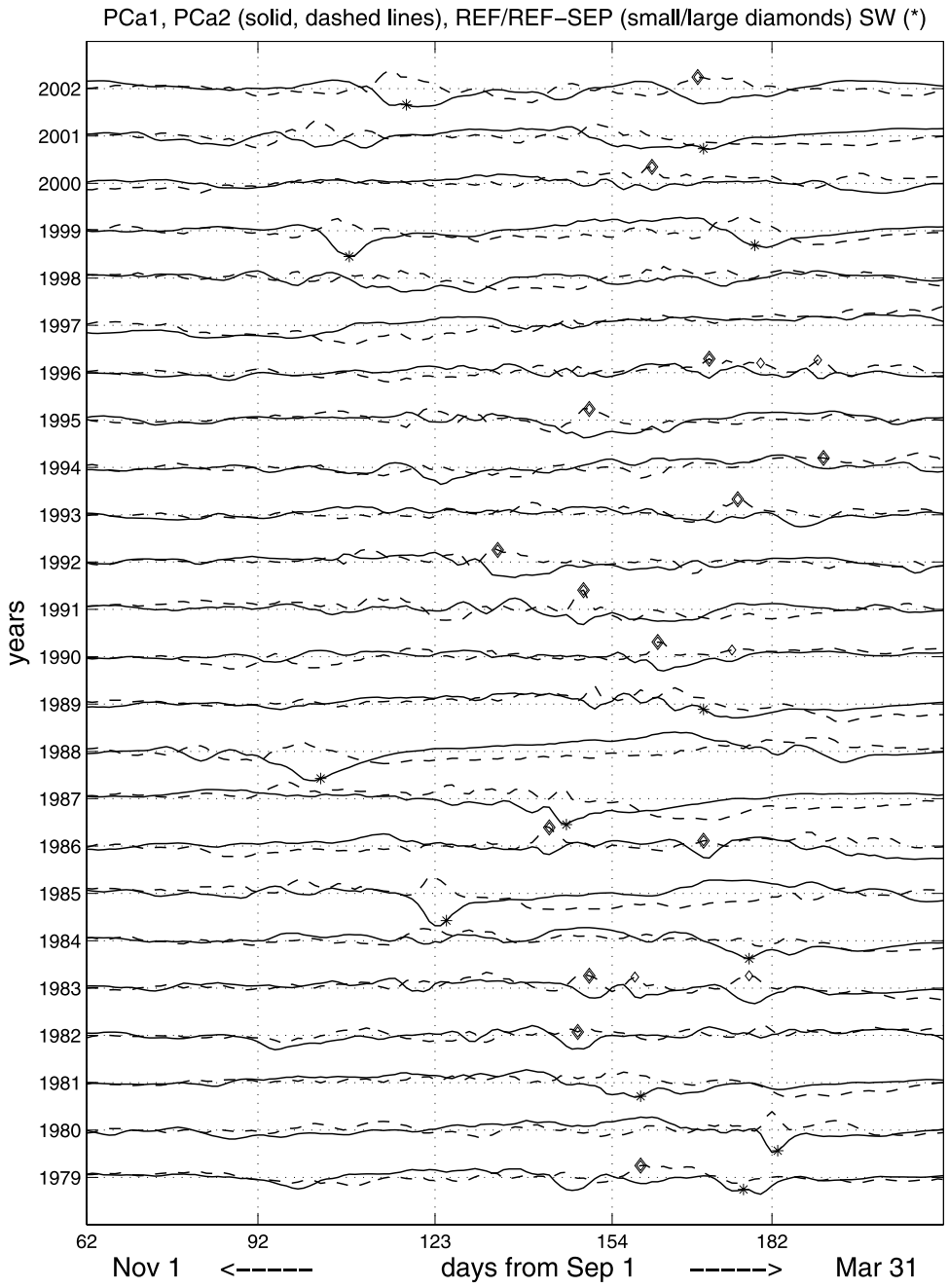

Figure A2. Daily plots of principal component time series (PC1, solid; PC2, dashed) of the EOFs of Figure A1, calculated from a singular value decomposition of the entire time series and plotted for each year individually. The central dates for SW events are marked on PC1 (stars), while the REF events are marked on PC2 (diamonds), with the REF-SEP events marked by larger diamonds.

Table A1. First Two EOF Statistics: Percent Variance, and the Correlations With U_{2-10}, U(10 hPa, 60 N), and U(10 hPa, 58–74 N) (See text for details)[a]

	Percentage of Variance [All Years, Yearly Mean (range)]	Correlations with U_{2-10}	Correlations with U(10 hPa, 60 N)	Correlations with U(10 hPa, 58–74 N)
EOF1	83	0.38	0.73	0.94
	[83 (69–92)]	[0.40 (−0.41–0.84)]	[0.71 (0.49–0.90)]	[0.94 (0.90–0.98)]
EOF2	14	−0.79	0.20	0.31
	[14 (6–27)]	[−0.76 (−0.97 – −0.54)]	[0.10 (−0.34 – 0.60)]	[0.22(−0.07 – 0.37)]

[a]EOFs are calculated using daily data from 1 September to 31 March, both using the entire data set, and using individual years. Results are presented as follows: (top) full data EOFs; (bottom in squared brackets) mean (min-max) of yearly EOFs.

variance and PC2 also accounting for a significant part. The results, along with correlation values, are also shown in Table A1.

[38] The above analysis indicates that two types of events, REF and SW, describe a large part of the high-latitude vortex variability. Indeed, a closer look at Figure A2 suggests that a large part of the negative PC1 events are accounted for by a SW event while many of the larger PC2 peaks are accounted for by a REF event. (If we define our REF events using a 1 standard deviation threshold (rather than 2), most of the positive PC2 peaks coincide with a REF event.) The REF and SW events obviously do not account for the opposite phases of the PCs, namely, a positive PC1, which corresponds to an overall vortex intensification [e.g., *Limpasuvan et al.*, 2005], and a negative PC2, which corresponds to a lower-mid stratospheric deceleration (*Abatzoglou and Magnusdottir* [2007] distinguished between upper level and lower mid-level wave breaking events). Note that while the entire 24-year PC1 and PC2 time series are not correlated (as should be in an EOF analysis), there can be nonzero correlations for individual years (as appears to be the case in Figure A2). The yearly EOF analysis yields yearly uncorrelated PC1 and PC2, but the time series do not correspond to REF and SW events as nicely, during some years. To gain more confidence in our analysis, we repeated it using shorter and longer time periods (winters with January 1991–2002 and 1958–2002). The resulting EOF structures as well as PC time series were similar to the analysis presented here (1979–2002).

Appendix B: The Holton-Mass Model

[39] Our model consists of solving a wave equation and a zonal mean equation, with the meridional structure of all fields projected onto sines and cosines, and truncated to a single meridional wave number (the channel width is half a meridional wave number). The setup is quite standard, and the reader is referred to the study by *Holton and Mass* [1976] (see also the study by *Plumb and Semeniuk* [2003]) for general details.

[40] Our channel is centered at 60°N, is about 10,000 km wide (equivalent to 90° latitude), the density-scale height is 7 km, the bottom model level is at 12.5 km (forcing is specified at 10 km) and the vertical grid is 2.5 km. The top is at 120 km (44 vertical grid points). Though our top lid is at 120 km, well within the large damping region, we also verified that we are not seeing downward reflection from the lid, by raising it to 240 km. This did not affect the results. The time step is 0.1 days, and we use a third-order Adams Bashford scheme for the time integration. The Brunt Vaisala frequency is $4 \cdot 10^{-4}$ sec^{-2} at all levels. We initiate the model with the zonally symmetric radiative equilibrium state (U_e, shown in Figure 6a), to which we relax the model with a timescale ranging from 5 days at upper levels, to 25 days at lower levels (also shown in Figure 6a). We keep the vertical shear at the top level equal to the radiative equilibrium value (which is zero in our runs), and keep the wind at the forcing level (10 km) constant with time. We specify the wave amplitude to be zero at the top level. At the bottom, we assume a constant phase (a stationary wave) and zonal wave number 1, and specify the wave amplitude to be zero at all times, except for a time period of $\pi/(\tau)$, during which it is a $\sin^2$:

$$\Phi_{bot}(5 < t < (\tau + 5)) = \Phi_o \sin^2\left(\frac{(t-5) * \pi}{\tau}\right)$$

where t and τ are in days. In our run series, we vary τ between 5–75 days in jumps of 10 days. In run series $R1$ we use Φ_o = 400 m, while in run series $R2$ Φ_{bot} varies between 404–104 m to keep $\int \Phi_{\mathrm{bot}}^2 dt$ at a fixed value. In run $R3$, we half Φ_{bot}, and increase the lowest level U_{re} from 10 to 20 m/s. We also ran the model with double vertical resolution, and with zonal wave number 2 and the results are similar.

Appendix C: The Index of Refraction and Vertical Wave Number Diagnostics

[41] The index of refraction relevant to the observations, for a stationary wave of zonal wave number s is:

$$N_{ref}^2 = \frac{a\overline{q}_\varphi}{U} - \frac{s^2}{\cos^2\varphi} + a^2 f^2 F(N^2) \qquad \text{(C1)}$$

where φ is latitude, a and f the earth's radius and Coriolis parameter, respectively, and $F(N^2)$ is a function of the Brunt Viasala frequency N^2, the exact form of which is given in the study by *Harnik and Lindzen* [2001].

[42] N_{ref}^2 is directly related to wave propagation in the meridional-height plane, through the QG wave equation:

$$\frac{a^2 f^2}{N^2}\frac{\partial^2 \phi}{\partial z^2} + \frac{f}{\cos\varphi}\frac{\partial}{\partial \varphi}\left(\cos\varphi \frac{\partial}{\partial \varphi}\left(\frac{\phi}{f}\right)\right) + N_{ref}^2 = damping \qquad \text{(C2)}$$

where z is log pressure height, and ϕ is the wave geopotential stream function, which is related to the actual wave geopotential height field Φ as follows: $\Phi = \phi e^{i(s\lambda)}\sqrt{\frac{N}{\rho}}$, where ρ is density and λ is longitude. The first two terms on the left hand side of equation (C2) represent vertical and meridional propagation, respectively. To relate the variations in the index of refraction to changes in wave structure, it is often necessary to separate the effects of the index of refraction on propagation in the vertical and meridional directions. This can be done diagnostically by dividing equation (C2) by $\frac{a^2 f^2}{N^2}\phi$, and taking the real part. The first term on the left hand side yields the vertical wave number squared $m_{(z)}^2$:

$$m_{(z)}^2 = -Re\left(\frac{\frac{\partial^2 \phi}{\partial z^2}}{\phi}\right) \qquad \text{(C3)}$$

[43] Correspondingly, the second term on the left hand side of equation (C2) yields a meridional wave number. To calculate $m_{(z)}^2$ for a given observed mean flow, we solve equation (C2) for ϕ, using the observed zonal mean U, $\frac{1}{a}\overline{q}_\varphi$, and N^2 fields, forced by a constant ϕ at the bottom level, and plug in equation (C3). For numerical purposes, we interpolate the observed fields to higher resolution, and

extend them vertically (to 105 km in this case), by keeping the fields constant. Extensive testing of the $m^2_{(z)}$ diagnostic shows that the main features are not sensitive to this interpolation/extrapolation. Details of the calculation are found in the study by *Harnik and Lindzen* [2001].

[44] In the Holton-Mass model, the meridional direction is separable from the vertical one, and it is commonly included in the definition of the refraction index, in a way similar to the zonal wave number contribution. For the Cartesian-coordinate β-plane model of Holton-Mass, this yields:

$$\tilde{N}^2_r = \frac{N^2}{f_o^2}\left(\frac{\overline{q}_y}{\gamma U} - k^2 - l^2 - \frac{f_o^2}{4H^2N^2}\right) \tag{C4}$$

where $\overline{q}_y = \beta + \gamma Q_y$, and $\gamma = \frac{8}{3\pi}$ is a meridional structure factor (from the projection of $\sin^2$ onto sin). We see that $\tilde{N}^2_r$ is similar to N^2_{ref}, except for an additional l^2 term. This term corresponds to the meridional propagation term in equation (C2) (second term on left hand). By including it in the definition of N^2_{ref}, we essentially get the 1-D version of $m^2_{(z)}$. Like $m^2_{(z)}$, a positive $\tilde{N}^2_r$ indicates vertical wave propagation, while a negative $\tilde{N}^2_r$ indicates vertical wave evanescence. Separating these regions are either a turning surface ($\tilde{N}^2_r = 0$) which reflects waves or by a critical surface ($U = 0$, $\tilde{N}^2_r \rightarrow \pm\infty$).

[45] **Acknowledgments.** The author thanks ECMWF for making the ERA-40 data available, Mark Cane for a useful suggestion regarding the Holton-Mass model, and three anonymous reviewers for very helpful comments. This work was supported by the Israeli Science Foundation grant 1370/08.

References

Abatzoglou, J., and G. Magnusdottir (2007), Wave breaking along the stratospheric polar vortex as seen in era-40 data, *Geophys. Res. Lett.*, *34*, L08812, doi:10.1029/2007GL029509.

Andrews, D. G., J. Holton, and C. B. Leovy (1987), *Middle Atmosphere Dynamics*, 489 pp., Academic Press, New York.

Baldwin, M. P., and T. J. Dunkerton (2001), Stratospheric harbingers of anomalous weather regimes, *Science*, *294*, 581–584.

Charlton, A., and L. Polvani (2007), A new look at stratospheric sudden warmings. part I: Climatology and modeling benchmarks, *J. Clim.*, *20*, 449–469.

Chen, P., and W. A. Robinson (1991), The effects of transience on the propagation of stratospheric planetary-waves, *J. Atmos. Sci.*, *48*, 1078–1092.

Chen, P., and W. A. Robinson (1992), Propagation of planetary-waves between the troposphere and stratosphere, *J. Atmos. Sci.*, *49*, 2533–2545.

Esler, J. G., and R. K. Scott (2005), On the excitation of transient Rossby waves on the polar stratospheric vortex and the barotropic sudden warming, *J. Atmos. Sci.*, *62*, 3661–3682.

Giannitsis, C., and R. S. Lindzen (2009), Nonlinear saturation of vertically propagating Rossby waves, *J. Atmos. Sci.*, *66*(4), 915–935.

Harnik, N., and E. Heifetz (2007), Relating over-reflection and wave geometry to the counter propagating Rossby wave perspective: Toward a deeper mechanistic understanding of shear instability, *J. Atmos. Sci.*, *64*, 2238–2261.

Harnik, N., and R. S. Lindzen (2001), The effect of reflecting surfaces on the vertical structure and variability of stratospheric planetary waves, *J. Atmos. Sci.*, *58*, 2872–2894.

Holton, J. R., and C. Mass (1976), Stratospheric vacillation cycles, *J. Atmos. Sci.*, *33*, 2218–2225.

Limpasuvan, V., D. W. J. Thompson, and D. L. Hartmann (2004), The life cycle of northern hemisphere sudden stratospheric warmings, *J. Clim.*, *17*, 2584–2596.

Limpasuvan, V., D. L. Hartmann, D. W. J. Thompson, K. Jeev, and Y. L.Yung (2005), Stratosphere-troposphere evolution during polar vortex intensification, *J. Geophys. Res.*, *110*, D24101, doi:10.1029/2005JD006302.

Lindzen, R. S. (1988), Instability of plane parallel shear-flow (toward a mechanistic picture of how it works), *Pure Appl. Geophys.*, *126*, 103–121.

McIntyre, M. E. (1982), How well do we understand the dynamics of stratospheric warmings, *J. Meteorol. Soc. Jpn.*, *46*, 37–65.

Perlwitz, J., and N. Harnik (2003), Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection, *J. Clim.*, *16*, 3011–3026.

Perlwitz, J., and N. Harnik (2004), Downward coupling between the stratosphere and troposphere: The relative roles of wave and zonal mean processes, *J. Clim.*, *17*, 4902–4909.

Plumb, R. A., and K. Semeniuk (2003), Downward migration of extratropical zonal wind anomalies, *J. Geophys. Res.*, *108*(D7), 4223, doi:10.1029/2002JD002773.

Scott, R., and L. Polvani (2004), Stratospheric control of upward wave flux near the tropopause, *Geophys. Res. Lett.*, *31*, L02115, doi:10.1029/2003GL017965.

Scott, R., and L. Polvani (2006), Internal variability of the winter stratosphere. part I: Time-independent forcing, *J. Atmos. Sci.*, *63*, 2758–2776.

Zhou, S., A. J. Miller, J. Wand, and J. K. Angell (2002), Downward-propagation temperature anomalies in the preconditioned polar stratosphere, *J. Clim.*, *15*, 781–792.

N. Harnik, Department of Geophysics and Planetary Sciences, Tel Aviv University, P.O. Box 39040, Kaplun 312, Ramat Aviv, Tel Aviv 69978, Israel. (harnik@tau.ac.il)